

- 1 -

SEQUENCE LISTING

<110> Autogen Research Pty Ltd

<120> Novel genes and their use in the modulation of obesity, diabetes and energy imbalance

<130> 2309315/TDO

<140>

<141>

<150> 60/141441

<151> 1999-06-29

<160> 22

<170> PatentIn Ver. 2.0

<210> 1

<211> 1669

<212> DNA

<213> mammalian

<220>

<221> CDS

<222> (43)..(1107)

<400> 1

ttcaaagagg	tgacatat	ttt ggttctg	atc cc	atg tgt	ttt tgg	54
				Met	Cys Phe Trp	

1

ggg ata	ttt ttg	tgt ttg	atc ttc	ctg gag	aaa agt	tgg gga	cag ata	102
Gly Ile	Phe Leu	Cys Leu	Ile Phe	Leu Glu	Lys Ser	Trp Gly	Gln Ile	
5		10		15		20		

caa atg	tcg tgt	tgg ccc	aag cct	ttg att	cca gaa	ctt gag	agg cag	150
Gln Met	Ser Cys	Trp Pro	Lys Pro	Leu Ile	Pro Glu	Leu Glu	Arg Gln	
	25		30		35			

aga tgc	acc gtt	gta aca	cca aaa	gtc ttc	cga gtc	gga gaa	tat gaa	198
Arg Cys	Thr Val	Val Thr	Pro Lys	Val Phe	Arg Val	Gly Glu	Tyr Glu	
	40		45		50			

caa gtt	aca ttt	gaa gcc	cac ggt	cac act	gac cca	ttt gat	gta acc	246
Gln Val	Thr Phe	Glu Ala	His Gly	His Thr	Asp Pro	Phe Asp	Val Thr	
	55		60		65			

- 2 -

atc tct ata aaa agt tac cct gat aaa aat gct aat tac tct tca agc	294
Ile Ser Ile Lys Ser Tyr Pro Asp Lys Asn Ala Asn Tyr Ser Ser Ser	
70 75 80	
tct gta cat tta tca cca gaa aat aaa ttc aaa aac tct aca atc tta	342
Ser Val His Leu Ser Pro Glu Asn Lys Phe Lys Asn Ser Thr Ile Leu	
85 90 95 100	
aca att cag ccc aaa cag ttg tct gaa ggg caa aac tcg tct tcg cat	390
Thr Ile Gln Pro Lys Gln Leu Ser Glu Gly Gln Asn Ser Ser Ser His	
105 110 115	
gtg tat ttg gaa gtt gtg tcc aag cat ttt tca aca tca aaa ata atg	438
Val Tyr Leu Glu Val Val Ser Lys His Phe Ser Thr Ser Lys Ile Met	
120 125 130	
tca atc gtc tat gac aat ggc act ctc ttc att cag act gac aag cct	486
Ser Ile Val Tyr Asp Asn Gly Thr Leu Phe Ile Gln Thr Asp Lys Pro	
135 140 145	
gtg tac act cca gag cag cct gta aag gtt gcc gtg tat tcg ctg gat	534
Val Tyr Thr Pro Glu Gln Pro Val Lys Val Ala Val Tyr Ser Leu Asp	
150 155 160	
gaa gcc tta aag cca gtc acc aga gag aca gtc tta acg ttc ata gac	582
Glu Ala Leu Lys Pro Val Thr Arg Glu Thr Val Leu Thr Phe Ile Asp	
165 170 175 180	
cct gaa gga tcc gaa gtt ggc ata gta gaa gga agc aat cat act gga	630
Pro Glu Gly Ser Glu Val Gly Ile Val Glu Gly Ser Asn His Thr Gly	
185 190 195	
atc acc tct ttc cct gac ttc agg att cct act aac cct aag ccc ggt	678
Ile Thr Ser Phe Pro Asp Phe Arg Ile Pro Thr Asn Pro Lys Pro Gly	
200 205 210	
aga tgg atg atc aag gct aaa tat aga gaa gat gct tca aca gct gga	726
Arg Trp Met Ile Lys Ala Lys Tyr Arg Glu Asp Ala Ser Thr Ala Gly	
215 220 225	
acc aca cac ttt gaa att aaa gag cat gat aaa gct ttc aaa ata gcc	774
Thr Thr His Phe Glu Ile Lys Glu His Asp Lys Ala Phe Lys Ile Ala	
230 235 240	
ctc gtt cca aca agt gat ctg gaa cac cca atg gaa gaa gca cgt ggc	822
Leu Val Pro Thr Ser Asp Leu Glu His Pro Met Glu Glu Ala Arg Gly	
245 250 255 260	
ctg agt ctc cag cca aaa aag tcc ctg caa gag atg ata cat gag caa	870
Leu Ser Leu Gln Pro Lys Lys Ser Leu Gln Glu Met Ile His Glu Gln	
265 270 275	
gct tcg aaa tac aaa cat cca gta ctg aag aaa tgt tgt tat gat gga	918

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

- 3 -

Ala Ser Lys Tyr Lys His Pro Val Leu Lys Lys Cys Cys Tyr Asp Gly
 280 285 290

gcc aga tat aac cac cat gaa acc tgt gag gaa cga gtt gcc cgt gtg 966
 Ala Arg Tyr Asn His His Glu Thr Cys Glu Glu Arg Val Ala Arg Val
 295 300 305

aaa ata ggc cca aac tgt gtc aga gcc ttc agt gaa tgc tgt gcc ctg 1014
 Lys Ile Gly Pro Asn Cys Val Arg Ala Phe Ser Glu Cys Cys Ala Leu
 310 315 320

gct agc gag aat acc ttt aag aat atc ctc atg tcg cgt ccc gat gac 1062
 Ala Ser Glu Asn Thr Phe Lys Asn Ile Leu Met Ser Arg Pro Asp Asp
 325 330 335 340

agt gga tat ttt act tta tct gct acc ata ctg gaa aat gct taa 1107
 Ser Gly Tyr Phe Thr Leu Ser Ala Thr Ile Leu Glu Asn Ala
 345 350 355

tcttattccc tgcaagtatt tgaagattac aagtattttc tgtgccttca cttttgctgg 1167

aaactaatgc acaaaatcaa acggagttca tacagcagtg aagcccttcc gctgtaactt 1227

tgccataaat agccttggtc gcacggaggt catttcataa ccgtaattta tccactggtc 1287

tcacaagtga gaccaagctg ataaaaacaa attcaccaga agagtttgat tgccatgcct 1347

agtgaccttg cccatcttcc tgtcaggacc ctcggtgccc taacatagta gaggggtgctc 1407

gggggacact caccgccaca aagaaagctg ccattccagcc ccggagagct gtggagtcaa 1467

cagcacacac cgtgtgggcc accgtgctgc ccaggtgtcc ataatgctac actaagtgcg 1527

cacgaataat cagttgtgcc agcagagtat gggagccgct aaaggatact atgcttgtaa 1587

atgtgtatca caatcagaat gtttaaataca ataaaatagt attgcccgcg ttaaaaaaaaa 1647

aaaaaaaaaa aaaaaaaaaa aa 1669

<210> 2
 <211> 354
 <212> PRT
 <213> mammalian

<400> 2
 Met Cys Phe Trp Gly Ile Phe Leu Cys Leu Ile Phe Leu Glu Lys Ser
 1 5 10 15

Trp Gly Gln Ile Gln Met Ser Cys Trp Pro Lys Pro Leu Ile Pro Glu
 20 25 30

Leu Glu Arg Gln Arg Cys Thr Val Val Thr Pro Lys Val Phe Arg Val

- 4 -

35	40	45
Gly Glu Tyr Glu Gln Val Thr Phe Glu Ala His Gly His Thr Asp Pro		
50	55	60
Phe Asp Val Thr Ile Ser Ile Lys Ser Tyr Pro Asp Lys Asn Ala Asn		
65	70	75 80
Tyr Ser Ser Ser Ser Val His Leu Ser Pro Glu Asn Lys Phe Lys Asn		
	85 90	95
Ser Thr Ile Leu Thr Ile Gln Pro Lys Gln Leu Ser Glu Gly Gln Asn		
100	105	110
Ser Ser Ser His Val Tyr Leu Glu Val Val Ser Lys His Phe Ser Thr		
115	120	125
Ser Lys Ile Met Ser Ile Val Tyr Asp Asn Gly Thr Leu Phe Ile Gln		
130	135	140
Thr Asp Lys Pro Val Tyr Thr Pro Glu Gln Pro Val Lys Val Ala Val		
145	150	155 160
Tyr Ser Leu Asp Glu Ala Leu Lys Pro Val Thr Arg Glu Thr Val Leu		
	165 170	175
Thr Phe Ile Asp Pro Glu Gly Ser Glu Val Gly Ile Val Glu Gly Ser		
	180 185	190
Asn His Thr Gly Ile Thr Ser Phe Pro Asp Phe Arg Ile Pro Thr Asn		
	195 200	205
Pro Lys Pro Gly Arg Trp Met Ile Lys Ala Lys Tyr Arg Glu Asp Ala		
210	215	220
Ser Thr Ala Gly Thr Thr His Phe Glu Ile Lys Glu His Asp Lys Ala		
225	230	235 240
Phe Lys Ile Ala Leu Val Pro Thr Ser Asp Leu Glu His Pro Met Glu		
	245 250	255
Glu Ala Arg Gly Leu Ser Leu Gln Pro Lys Lys Ser Leu Gln Glu Met		
	260 265	270
Ile His Glu Gln Ala Ser Lys Tyr Lys His Pro Val Leu Lys Lys Cys		
275	280	285
Cys Tyr Asp Gly Ala Arg Tyr Asn His His Glu Thr Cys Glu Glu Arg		
290	295	300
Val Ala Arg Val Lys Ile Gly Pro Asn Cys Val Arg Ala Phe Ser Glu		
305	310	315 320

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

- 5 -

Cys Cys Ala Leu Ala Ser Glu Asn Thr Phe Lys Asn Ile Leu Met Ser
 325 330 335

Arg Pro Asp Asp Ser Gly Tyr Phe Thr Leu Ser Ala Thr Ile Leu Glu
 340 345 350

Asn Ala

<210> 3

<211> 1170

<212> DNA

<213> mammalian

<220>

<221> CDS

<222> (20)..(589)

<400> 3

```

gtcgttggtt tcggcggcc atg gag agc gca gag gag cct ctg ccc gcg cgg   52
           Met Glu Ser Ala Glu Glu Pro Leu Pro Ala Arg
                1                 5                 10

ccg gcg ctg gag acc gag ggc ctg agg ttc ctg cac gtc aca gtg ggc   100
Pro Ala Leu Glu Thr Glu Gly Leu Arg Phe Leu His Val Thr Val Gly
                15                 20                 25

tcc ctg ctg gcc agc tat ggc tgg tac gtc ctc ttc agc tgc atc ctt   148
Ser Leu Leu Ala Ser Tyr Gly Trp Tyr Val Leu Phe Ser Cys Ile Leu
                30                 35                 40

ctc tac att gtc atc cag aag ctc tcc gtc cga ttg agg gtt ttg agg   196
Leu Tyr Ile Val Ile Gln Lys Leu Ser Val Arg Leu Arg Val Leu Arg
                45                 50                 55

cag agg cag ctg gac cag gct gac gct gtt ctg gaa cct gat gct gtt   244
Gln Arg Gln Leu Asp Gln Ala Asp Ala Val Leu Glu Pro Asp Ala Val
                60                 65                 70                 75

gtt aag cga caa gag gct tta gcc gct gct cgt ttg aga atg cag gaa   292
Val Lys Arg Gln Glu Ala Leu Ala Ala Ala Arg Leu Arg Met Gln Glu
                80                 85                 90

```

- 6 -

gat cta aat gcc caa gtt gaa aag cat aag gaa aaa cta aga cag ctt 340
Asp Leu Asn Ala Gln Val Glu Lys His Lys Glu Lys Leu Arg Gln Leu
95 100 105

gaa gaa gaa aaa agg aga cag aag att gaa atg tgg gac agc atg caa 388
Glu Glu Glu Lys Arg Arg Gln Lys Ile Glu Met Trp Asp Ser Met Gln
110 115 120

gaa ggc aga agt tac aga aga aat cca gga agg cct cag gaa gaa gat 436
Glu Gly Arg Ser Tyr Arg Arg Asn Pro Gly Arg Pro Gln Glu Glu Asp
125 130 135

ggt cct gga cct tct act tca tca tct gtc acc cgc aaa gga aaa tct 484
Gly Pro Gly Pro Ser Thr Ser Ser Ser Val Thr Arg Lys Gly Lys Ser
140 145 150 155

gac aaa aag cct ttg agg gga aat ggt tat aac cct ctg acg ggt gaa 532
Asp Lys Lys Pro Leu Arg Gly Asn Gly Tyr Asn Pro Leu Thr Gly Glu
160 165 170

ggg ggt gga acc tgc gcc tgg aga cct gga cgc agg ggc cca tca tct 580
Gly Gly Gly Thr Cys Ala Trp Arg Pro Gly Arg Arg Gly Pro Ser Ser
175 180 185

ggt gga tga agctaagacc cttgttagtg tcgctttgac attagcaagg. 629
Gly Gly
190

tgaaccctta accctcaact cagttgcctt acgcacactt tcacagtgac tagccaagga 689

gaggtggggc ttatttccat tcgtagctac ctgtattcta agggctttgg tcagtgtgag 749

ctatggacat tgtcattagg tcatattcta cttagacaac agtcattgat ttcattggcta 809

cttgctagtt gatagggttaa aggcctctcg ctgttttagca aacttcataa aggaggccca 869

gtgatgatcc tttggggtag aagtccttgc tgacaggatg gtctctgtga caggatgcgt 929

tcaatgatgt cttccttata aatgggtgagc ccaccagtga ggattactga tgtgcacagt 989

tgatgggggtt tgcttctgta tatttatttt tatgtacaga aatttgcaaa aaaaaataaa 1049

aagtaacatt tttagcatct ttattaaact caaggaaatt tcgttgtgag cttgactttg 1109

tctatcagac attaaacagc tttttatcat taaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1169

a 1170

- 7 -

<210> 4

<211> 189

<212> PRT

<213> mammalian

<400> 4

Met Glu Ser Ala Glu Glu Pro Leu Pro Ala Arg Pro Ala Leu Glu Thr
 1 5 10 15

Glu Gly Leu Arg Phe Leu His Val Thr Val Gly Ser Leu Leu Ala Ser
 20 25 30

Tyr Gly Trp Tyr Val Leu Phe Ser Cys Ile Leu Leu Tyr Ile Val Ile
 35 40 45

Gln Lys Leu Ser Val Arg Leu Arg Val Leu Arg Gln Arg Gln Leu Asp
 50 55 60

Gln Ala Asp Ala Val Leu Glu Pro Asp Ala Val Val Lys Arg Gln Glu
 65 70 75 80

Ala Leu Ala Ala Ala Arg Leu Arg Met Gln Glu Asp Leu Asn Ala Gln
 85 90 95

Val Glu Lys His Lys Glu Lys Leu Arg Gln Leu Glu Glu Glu Lys Arg
 100 105 110

Arg Gln Lys Ile Glu Met Trp Asp Ser Met Gln Glu Gly Arg Ser Tyr
 115 120 125

Arg Arg Asn Pro Gly Arg Pro Gln Glu Glu Asp Gly Pro Gly Pro Ser
 130 135 140

Thr Ser Ser Ser Val Thr Arg Lys Gly Lys Ser Asp Lys Lys Pro Leu
 145 150 155 160

Arg Gly Asn Gly Tyr Asn Pro Leu Thr Gly Glu Gly Gly Gly Thr Cys
 165 170 175

Ala Trp Arg Pro Gly Arg Arg Gly Pro Ser Ser Gly Gly
 180 185

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

- 8 -

```
<210> 5
<211> 1174
<212> DNA
<213> mammalian
```

```
<220>  
<221> CDS  
<222> (31) .. (594)
```

<400>	5															
cagggctggg	gggcggcggc	ggcggcggtc	atg	gaa	cgc	caa	gag	gag	tct	ctg	54					
			Met	Glu	Arg	Gln	Glu	Glu	Ser	Leu						
			1				5									
tcc	gcg	cgg	ccg	gcc	ctg	gag	acc	gag	ggg	ctg	cgc	ttc	ctg	cac	acc	102
Ser	Ala	Arg	Pro	Ala	Leu	Glu	Thr	Glu	Gly	Leu	Arg	Phe	Leu	His	Thr	
	10					15					20					
acg	gtg	ggc	tcc	ctg	ctg	gcc	acc	tat	ggc	tgg	tac	atc	gtc	ttc	agc	150
Thr	Val	Gly	Ser	Leu	Leu	Ala	Thr	Tyr	Gly	Trp	Tyr	Ile	Val	Phe	Ser	
	25				30					35					40	
tgc	atc	ctt	ctc	tac	gtg	gtc	ttt	cag	aag	ctt	tcc	gcc	cgg	cta	aga	198
Cys	Ile	Leu	Leu	Tyr	Val	Val	Phe	Gln	Lys	Leu	Ser	Ala	Arg	Leu	Arg	
				45					50						55	
gcc	ttg	agg	cag	agg	cag	ctg	gac	cga	gct	gcg	gct	gct	gtg	gaa	cct	246
Ala	Leu	Arg	Gln	Arg	Gln	Leu	Asp	Arg	Ala	Ala	Ala	Ala	Val	Glu	Pro	
			60					65					70			
gat	gtt	gtt	gtt	aaa	cga	caa	gaa	gct	tta	gca	gct	gct	cga	ctg	aaa	294
Asp	Val	Val	Val	Lys	Arg	Gln	Glu	Ala	Leu	Ala	Ala	Ala	Arg	Leu	Lys	
		75					80					85				
atg	caa	gaa	gaa	cta	aat	gcg	caa	gtt	gaa	aag	cat	aag	gaa	aaa	ctg	342
Met	Gln	Glu	Glu	Leu	Asn	Ala	Gln	Val	Glu	Lys	His	Lys	Glu	Lys	Leu	
	90					95					100					
aaa	caa	ctt	gaa	gaa	gaa	aaa	agg	aga	cag	aag	att	gaa	atg	tgg	gac	390
Lys	Gln	Leu	Glu	Glu	Glu	Lys	Arg	Arg	Gln	Lys	Ile	Glu	Met	Trp	Asp	
105					110					115					120	
agc	atg	caa	gaa	gga	aaa	agt	tac	aaa	gga	aat	gca	aag	aag	ccc	cag	438
Ser	Met	Gln	Glu	Gly	Lys	Ser	Tyr	Lys	Gly	Asn	Ala	Lys	Lys	Pro	Gln	
				125					130					135		
gag	gaa	gac	agt	cct	ggg	cct	tcc	act	tca	tct	gtc	ctg	aaa	cgg	aaa	486
Glu	Glu	Asp	Ser	Pro	Gly	Pro	Ser	Thr	Ser	Ser	Val	Leu	Lys	Arg	Lys	
			140					145					150			
ccg	gac	aga	aag	cct	ttg	cgg	gga	gga	ggt	tat	aac	ccg	ttg	tct	ggt	534
Ser	Asp	Arg	Lys	Pro	Leu	Arg	Gly	Gly	Gly	Tyr	Asn	Pro	Leu	Ser	Gly	
	155						160					165				

- 9 -

gaa gga ggc gga gct tgc tcc tgg aga cct gga cgc aga ggc ccg tca 582
 Glu Gly Gly Gly Ala Cys Ser Trp Arg Pro Gly Arg Arg Gly Pro Ser
 170 175 180

tct ggc gga tga ggctaagaat cttgttagtg tcaacttttga cattagcaag 634
 Ser Gly Gly
 185

atgaaccctt aaccctcgat tcaattgcct tacgcacgct tttcacagtg actagccaag 694
 gggaggtggg gttgatttct gttcctaact acacctgcat atgtcagggc tccagtcagc 754
 aaaaggtata gatgttgccct ctaggcatga ggtcattggg cacattctac ttggagacag 814
 tgattgcatt cattgatttc atgggtaatt gctagttggg aggttaaaggc ctctagatga 874
 ttagcaatct tgataaaaaga ggcctagtaa tgttcttttg aggttagaaa tccttgctgc 934
 taggacagtc tctgtgacag gttgcgttga atgatgtctt ccttatcaat ggtgagccca 994
 ccagtgagga ttactgatgt ggacagttga tgggggtttgt ttctgtatat ttatttttat 1054
 gtacagaact ttgtaaaaaac gaaactatct aaaaaacaag aataacattt ttagcatctt 1114
 tattcaagga gatttatgga cttcaatttg tctatcaaac attaaatagc tttttattac 1174

<210> 6
 <211> 187
 <212> PRT
 <213> mammalian

<400> 6
 Met Glu Arg Gln Glu Glu Ser Leu Ser Ala Arg Pro Ala Leu Glu Thr
 1 5 10 15
 Glu Gly Leu Arg Phe Leu His Thr Thr Val Gly Ser Leu Leu Ala Thr
 20 25 30
 Tyr Gly Trp Tyr Ile Val Phe Ser Cys Ile Leu Leu Tyr Val Val Phe
 35 40 45
 Gln Lys Leu Ser Ala Arg Leu Arg Ala Leu Arg Gln Arg Gln Leu Asp
 50 55 60
 Arg Ala Ala Ala Ala Val Glu Pro Asp Val Val Val Lys Arg Gln Glu
 65 70 75 80
 Ala Leu Ala Ala Ala Arg Leu Lys Met Gln Glu Glu Leu Asn Ala Gln
 85 90 95
 Val Glu Lys His Lys Glu Lys Leu Lys Gln Leu Glu Glu Glu Lys Arg

- 10 -

100	105	110
Arg Gln Lys Ile Glu Met Trp Asp Ser Met Gln Glu Gly Lys Ser Tyr		
115	120	125
Lys Gly Asn Ala Lys Lys Pro Gln Glu Glu Asp Ser Pro Gly Pro Ser		
130	135	140
Thr Ser Ser Val Leu Lys Arg Lys Ser Asp Arg Lys Pro Leu Arg Gly		
145	150	155
Gly Gly Tyr Asn Pro Leu Ser Gly Glu Gly Gly Gly Ala Cys Ser Trp		
165	170	175
Arg Pro Gly Arg Arg Gly Pro Ser Ser Gly Gly		
180	185	

<210> 7
 <211> 279
 <212> DNA
 <213> mammalian

<220>
 <221> CDS
 <222> (54)..(140)

<400> 7
 ctgaaaaggc tgttgtcaag atggagtgtc taaccagta atccaaggac caa atg 56
 Met
 1
 ctg agt cca cac agt gtg gcc agc atg ctg tct gca gtt gaa gca ggg 104
 Leu Ser Pro His Ser Val Ala Ser Met Leu Ser Ala Val Glu Ala Gly
 5 10 15
 aca gtt ttt ctt cta gtg act agc tta cca cat tga ggcaaactcc 150
 Thr Val Phe Leu Leu Val Thr Ser Leu Pro His
 20 25
 atgtggaggt tcttcgatgc tcatcatctt ctttgaagtg gagtggagac gctgccagaa 210
 gcagacgtgt ttcactgggc aagaaagcct tttattaata aaacatctca aatgccataa 270
 aaaaaaaaaa 279

<210> 8
 <211> 28
 <212> PRT
 <213> mammalian

<400> 8

- 11 -

Met Leu Ser Pro His Ser Val Ala Ser Met Leu Ser Ala Val Glu Ala
 1 5 10 15

Gly Thr Val Phe Leu Leu Val Thr Ser Leu Pro His
 20 25

<210> 9

<211> 5251

<212> DNA

<213> mammalian

<400> 9

cagggctggg cggcggcggc ggcggcggtc atggaacgcc aagaggagtc tctgtccgcg 60
 cggcggcccc tggagaccga ggggctgcgc ttctgcaca ccacgggtga gtcgttgccg 120
 ggcagccggg cgcgcgcgc cacttttgcg acgcgcagcc atgatgggtg ggtcgtccgc 180
 cgtgcaccg ggcgcggag cctgggaggc ctgggaacgg tcgggcgttg gcgcttacgc 240
 ggaccttggg cagcaggccc ggaccttgcg cggaggcttc tcgggagccg cacttccctg 300
 ggcggctcgg ctgtcccttg tttgcgcaag tcttttttgc gaaccaagcc ctctctgtgg 360
 tagttactgg ggtcactcgg cgttgggcgt ttgcctctgg gaccgcgcc acacagcccc 420
 atacacactc ctgactcccc gcgctgtcac ccctttctat gtggctctga aaggcctttg 480
 ccttctgat tcagattagt tgctcttcac tcttcaaaac ccagttgctg tgccctccac 540
 actctaactg ccccgactc ccagatggt tgggaagtct cacttctcag tgatccctga 600
 attgtcgcac ttcttgagtt cgtgttttaa cgatctactt aggaggcttt ttctcagcc 660
 tagaccatga aggctttgag ggcaggagtt acactttgtg tttgttgagt cttatggaaa 720
 ggtcaactag tagtgtcatt tttagttttt tgaaaactgt ttttcttttc agtgggctcc 780
 ctgctggcca cctatggctg gtacatcgtc ttcagctgca tccttctcta cgtggctctt 840
 cagaagcttt ccgcccggct aagagccttg aggcagaggc agctggaccg agctgcggct 900
 gctgtgggtt agtgccctgat aaccgaaatg aaagcgggtg ttttgacct cctttatatt 960
 aagagttagt ctcttagtaa aagtaagagg ggccacacag gaagacctg tctctattta 1020
 aaaaaaaaaa aaatagccgg gagtggcggc acgcacctgt agtccagct gctcaggagg 1080
 ctgaggcggg ataatcactt gagtccaggg agtcaaagct gcagtgggct atgctcgggc 1140
 cacactacac tccagcctgg gcaattgatt gagacctgtg ctttaaaaaa aaaaaaaaaa 1200

- 12 -

aaaaagtagg aagtatatgg ttctcgggtgg ggcgcggtgg ctcacacctg taatcccagc 1260
 actttgggaa gccgaggcag gaggatgact tgaggtcagg ggttcgagaa cagcctggcc 1320
 aacatggtga aaccctgtct ctactaaaaa tacaaatatt agtggggcgt ggtgacgggc 1380
 acctgtaatc ccagctatta ggggtggctga ggcaggagaa atcgcttgaa cctgggagct 1440
 ggagattgca gtgagctgag attgtgccac tgcactccag cctgggcaac agagtggagac 1500
 tgtcttttct ttcttttttt tttttttttc tatgagatgg agtctagcct tgttgcaaag 1560
 agcgagactc tatgagtaga cgttatgaat agaaatgagt tcattttctat tcataatgct 1620
 atttggaagg atttttcttt tctgtagaaa caaatactta agaatcttct gcgctaatta 1680
 agggatggat aatgatttag aaaactttat atttccttgg tagtcttcca ggattctagt 1740
 cagcctagag actgtgggtg tcaactgaggt atccaagatg tgctctgtgt ggccactatc 1800
 ccaggcttta tgaatcggaa ttgtcaggg gaactcagaa attggcattt ctaacagatt 1860
 tctggtgatg tagatatttc gggctaaaaa ccgtggctca gcaacagacc cctgccccct 1920
 gaagcagtaa aatgtatgca gagggggttag gagtacttat gtaaaaaatat gttgtttcat 1980
 tgtctgatat ccataacctt ttatactttt aataatatgg acactcaaaa gtttctattt 2040
 tatattgtac acagtgtttt atctccattt ttttctgaca ttttagaacc tgatgttggt 2100
 gttaaacgac aagaagcttt agcagctgct cgactgaaaa tgcaagaaga actaaatgct 2160
 caagttgaaa agcataagga aaaactgaaa caagtatgaa ctgggtttcag tttgaatgtg 2220
 tgcatagaaa ttgtctgagg tttagtggct aacgatgcct gtgtctgtgt tgtctataag 2280
 cttctaggac caggctctat ccattagat tcaataagca tttcagttcc taccatgtaa 2340
 gtattggtga tatcaagaag aatacacgat tgtaggggaa cactagatgt gtgaatatat 2400
 taccatgaaa ggtccagagc acaaaaggag ggacaggctg gagcaggag catgtgagtg 2460
 tgtgtgtgca tgtgcctgtg tcttcccat taccaaaaat gtcctgacag gagtgagttt 2520
 cagaagaatg gagtcagtaa tctttttcat gaaacatttt gctttcttta atagtgtaca 2580
 aaaaccaaag ctgctctatg tgagttaaac tcacactacc agatcacaac agttttatta 2640
 actaaagaaa acgagggtga agtttgttct gaaagacatt taaattaaga attatcagag 2700
 ttagctttgt ctttgagaga aatggcagct tctgaattct ttctgtaaaa tgtgattgtt 2760
 totcagcttg aagaagaaaa aaggagacag aagattgaaa tgtgggacag catgcaagaa 2820

- 13 -

ggaaaaagtt acaaaggaaa tgcaaagaag cccaggtga ctggagacct cggccggctg 2880
 gcatgcggtg gatgaagatt gccaaagtaga atgttttaat tgcttcttac actactgtgt 2940
 gtgttcaaac aggaggaaga cagtcctggg ccttccactt catctgtcct gaaacggaaa 3000
 tcggacagaa agcctttgcg gggaggaggt aagcaccact gatgtcaaat gttaacagat 3060
 tttcaacact tacaggatat agttacctt taagaacaag attgtttgtt tctttgtcca 3120
 taaattaaga ctaattcctt aggattgtga agattcaata aaggaaacag atgcaaatca 3180
 cctcctaggt cctcactaag tacttagaag gattgtactt atagtattct aacttgatcc 3240
 ttctgcagcc ccgtagaggg agagctaagt aggggtgagga attgtctgcc aatcttcaga 3300
 tgagtgtcaa ggagctggaa cacagtgggt ttgggtctttc tggtctgggac caccttgttt 3360
 cttgcaaata acaaggagta gcagacagat gctcatccaa agctgcttcc tgtgtgcagc 3420
 actgccccgg ggactctgga tgatgccaca gcagtctgtc ttcattccat ccttgagaat 3480
 ttcaaactctg ggaagatggg actcacaac gaaaataagc aatccttggg gattctggct 3540
 aagagttgca agttactgct gaggaaggaa agaacaaca cactagaaca ctgtaggaac 3600
 caaggcgga gattttgtat cctccatagg aggagagggg caccgcagag gccctgatgg 3660
 tgtctttgag gactgaggaa agactggggc atgggctcca aggcagcagg gccacagact 3720
 tggctgacct taaacgctga gctgtaatcc cctttgtgtc agaagactaa acctggcttg 3780
 ctgtagagaa ggtgatgcat ctggaaagaa aatgctattt ttaaattggc ctgccggaag 3840
 cttattttta gacacataga ggtgatattt aggagaggaa tggaaatcgt agaagatgga 3900
 atgcagggtg tgcttgctg cacggcctct ttcagcatcc ccagcatttc tgagctggga 3960
 cttttgacta gcctggcttt acaataagg aaactgaggc acagtgttta attgccaaa 4020
 gattccacta taagtaagga gtaaaagtaa catttaagtt ctgggtggcc ctagaacctt 4080
 agcactcaac caggttacca gttgtgcact gactttggga agctcatgag ggagtggggg 4140
 ggttgggggt agggaaggat acagaagacc ccgttctgac tggtagaagt gacaagtttg 4200
 actcttgatt ttttttaatc tgttttctgt agcgtgaaca gcccttattt gaatgtatga 4260
 gttttagtaa gcactgtgat aggaggattc atatacttaa atcaggccct cttgagagag 4320
 ttttttggtg acccttttgc atgtgtttcg gaggttggga caaagaagct gaatgacttt 4380

- 14 -

```

tttccccacc agacaatcag ttcaaattggc aatcacacaata taaagggtttt ttttttttttc 4440
acatagctaa aagggtttttt taaatgtccc ttaggatctg tatcttttgca gtgcttttgcg 4500
tgtcactctc ataattttat tgtggatata caatgttccc agatttttcag atttttatca 4560
atactgttgt gctgcttttc tgtcctccca ggttataacc cgttgtctgg tgaaggaggc 4620
ggagcttgct cctggagacc tggacgcaga ggcccgatcat ctggcggatg aggctaagaa 4680
tcttgtagt gtcacttttg acattagcaa gatgaacct taacctcga ttcaattgcc 4740
ttacgcacgc ttttcacagt gactagccaa ggggagggtg ggttgatttc tgttcctaac 4800
tacacctgca tatgtcaggg ctccagtcag caaaaggat agatgttgcc tctaggcatg 4860
aggtcattgg tcacattcta cttggagaca gtgattgcat tcattgattt catgggtaat 4920
tgctagttag taggtaaagg cctctagatg attagcaatc ttgataaaaag aggcctagta 4980
atgttctttt gaggttagaa atccttgctg ctaggacagt ctctgtgaca ggttgcggtg 5040
aatgatgtct tccttatcaa tgggtgagccc accagtgagg attactgatg tggacagttg 5100
atgggggttg tttctgtata tttattttta tgtacagaac tttgtaaaaa cgaaactatt 5160
taaaaaacaa gaataacatt tttagcatct ttattcaagg agatttatgg acttcaattt 5220
gtctatcaaa cattaaatag ctttttatta c 5251

```

```

<210> 10
<211> 21
<212> DNA
<213> mammalian

```

```

<400> 10
gggagagctg tggagtcaac a 21

```

```

<210> 11
<211> 22
<212> DNA
<213> mammalian

```

```

<400> 11
cgtggcgact tagtgtagca tt 22

```

```

<210> 12
<211> 23
<212> DNA
<213> mammalian

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

- 15 -

<400> 12
gatgcgttca atgatgtctt cct 23

<210> 13
<211> 22
<212> DNA
<213> mammalian

<400> 13
agaagcaaac cccatcaact gt 22

<210> 14
<211> 21
<212> DNA
<213> mammalian

<400> 14
tggaggttct tcgatgctca t 21

<210> 15
<211> 22
<212> DNA
<213> mammalian

<400> 15
cagtgaaaca cgtctgcttc tg 22

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144

- 16 -

<210> 16

<211> 22

<212> DNA

<213> mammalian

<400> 16

gcaaagacct gtatgccaac ac

22

<210> 17

<211> 23

<212> DNA

<213> mammalian

<400> 17

gccagagcag tgatctcttt ctg

23

<210> 18

<211> 21

<212> DNA

<213> mammalian

<400> 18

accgtgctgc ccagggtgcc a

21

<210> 19

<211> 29

<212> DNA

<213> mammalian

<400> 19

tgagcccacc agtgaggatt actgatgtg

29

<210> 20

<211> 29

<212> DNA

<213> mammalian

<400> 20

atcttctttt aagtggagtg gagacgctg

29

<210> 21

<211> 25

<212> DNA

<213> mammalian

<400> 21

tccggtccac aatgcctggg tacat

25

- 17 -

<210> 22

<211> 18

<212> PRT

<213> mammalian

<400> 22

Arg	Pro	Gln	Glu	Glu	Asp	Gly	Pro	Gly	Pro	Ser	Thr	Ser	Ser	Ser	Val
1				5					10					15	

Thr Arg

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000
 1001
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1